

Similarity-Projection structures: the logical geometry of Quantum Physics *

Daniel Lehmann

School of Engineering and Center for the Study of Rationality
Hebrew University,
Jerusalem 91904, Israel

May 2008

Abstract

Similarity-Projection structures abstract the numerical properties of real scalar product of rays and projections in Hilbert spaces to provide a more general framework for Quantum Physics. They are characterized by properties that possess direct physical meaning. They provide a formal framework that subsumes both classical boolean logic concerned with sets and subsets and quantum logic concerned with Hilbert space, closed subspaces and projections. They shed light on the role of the phase factors that are central to Quantum Physics. The generalization of the notion of a self-adjoint operator to SP-structures provides a novel notion that is free of linear algebra. Keywords: Similarity-Projection structures, Measurement algebras, Quantum Logic. PACS: 02.10.-v.

1 Introduction

In [5], H. Whitney abstracted the properties of linear dependence from the setting of vector spaces. This paper represents a similar endeavor to abstract

*This work was partially supported by the Jean and Helene Alfassa fund for research in Artificial Intelligence

the properties of both linear dependence and projections on closed subspaces from the vector space structure of Hilbert spaces.

A system of Quantum Physics is described by a set Ω of pure states. In traditional presentations those pure states are modeled as rays, i.e., one-dimensional subspaces, of a Hilbert space. The main structure possessed by Ω is its real scalar product. Given any two rays x, y their real scalar product $p(x, y)$ is a real number in the interval $[0, 1]$, customarily described as a *transition probability*. This quantity is physically meaningful and can be measured in experiments. It seems to be the only physically meaningful quantity: the only physical property that can be directly measured. The purpose of this paper is to study the properties of this quantity.

This paper studies the properties of the real scalar product of rays in Hilbert spaces. Surprisingly, such a study has not been pursued very actively so far. An algebraic characterization of the properties of real scalar product of rays in Hilbert space would be interesting, but is not the primary goal we are seeking. Some of those properties are not satisfied by the spaces in which Quantum Physics is done, which include superselection rules. For example the following is a property of p , the real scalar product of rays that is satisfied in all Hilbert spaces but not when superselection rules are introduced: for any distinct rays x, y , there exists a ray z such that $0 < p(x, z) < 1$. This paper's goal is to propose a list, as extensive as possible, of properties of the real scalar product of rays *that are physically meaningful and satisfied in all spaces used by Quantum Physics, including classical systems and superselection rules*.

Phase factors play a central role in the thinking of Quantum physicists. We shall examine the nature of those phase factors, ask whether they can be defined in terms of the real scalar product of rays. We shall see that the cosine of those phase factors are definable in terms of the real scalar product of rays. We shall consider whether the phase factors themselves have physical meaning or whether only some trigonometric function of those phase factors is meaningful.

In a previous paper Lehmann, Engesser and Gabbay in [3] proposed a qualitative study of projections in Hilbert spaces and proposed M-algebras as an abstraction of properties of projections in Hilbert spaces meaningful for Quantum Physics. The present paper builds on this first effort and shares its philosophy. This paper is a direct successor of [2] which is a concrete study of Hilbert spaces, but failed to give a proper analysis of the phase factors and of [1] which proposes a numberless analysis of projections on subspaces.

2 Similarity

The Similarity-Projection structures (from now on, SP-structures) will be introduced gently and slowly. We are defining structures that include a non-empty set (the carrier) Ω . The elements of Ω are to be thought of as pure states. Elements of Ω will indeed be called states.

A characteristics of Quantum Physics is that pure states have a dual aspect: they are both states and questions (i.e, observables). A state $x \in \Omega$ can be understood both as a state as in “the system is in state x ” and as a question like “let us measure whether the system is in state x or not”. Given two states x and y , if a system in state x is asked whether it is in y , there is, in Quantum Physics, a certain “probability” that the answer will be positive. Given two pure states s_1 and s_2 , one can measure the probability that one will obtain s_2 when measuring, in state s_1 , whether s_2 holds or not. Think, for example, about the simplest of quantum systems: a particle of spin $1/2$, let s_1 be the state $|+\rangle$ in which the spin is up in the z -direction, and s_2 be the state in which the spin is up in the x -direction. We know that the measurement of the spin in the x -direction will, on a system in state s_1 give the answer *up* with probability $1/2$ and the answer *down* with the same probability.

The first structural ingredient in the definition of SP-structures is therefore a real function $p : \Omega \times \Omega \longrightarrow \mathcal{R}$. If x and y are states, the real number $p(x, y)$ is to be understood as the *similarity* of x to y , or, in the language used by physicists, the *transition probability* between x and y .

3 Hilbert and classical SP-structures

We shall now present two paradigmatical examples of such similarity functions p . The first example covers what we shall call Hilbert SP-structures. Assume \mathcal{H} is a Hilbert space and Ω is the set of unit vectors of \mathcal{H} . For any $\vec{x}, \vec{y} \in \Omega$ define $p(\vec{x}, \vec{y})$ to be the real scalar product of \vec{x} and \vec{y} : $p(\vec{x}, \vec{y}) = |\langle \vec{x}, \vec{y} \rangle|^2$. Note that we depart from the presentation that dates back at least to von Neumann of taking Ω to be the set of rays, i.e., one-dimensional subspaces of \mathcal{H} . We consider unit vectors, not rays. This is, in fact, closer to the every day practice of physicists.

The second example consists of an arbitrary set Ω and a similarity function defined by: $p(x, y) = 1$ if $x = y$ and $p(x, y) = 0$ otherwise. We shall call

such structures *classical* SP-structures.

4 Symmetry and Non-negativity

We shall now list a number of properties of the similarity p that we want to assume in any SP-structure. We shall draw some consequences of those assumptions as we proceed.

Since we are dealing with structures of the type $\langle \Omega, p \rangle$ it is natural to define as equivalent any two elements of Ω that behave in exactly the same way as far as p is concerned.

Definition 1 *Any two states $x, y \in \Omega$ are said to be equivalent, and we write $x \sim y$ iff for any $z \in \Omega$, one has: $p(x, z) = p(y, z)$.*

The relation \sim is obviously an equivalence relation. In classical SP-structures, one has $x \sim y$ iff $x = y$. In Hilbert SP-structures two unit vectors are equivalent iff they differ by a phase factor.

4.1 Symmetry

Our first assumption is a symmetry assumption.

Property 1 (Symmetry) *For any $x, y \in \Omega$, $p(y, x) = p(x, y)$.*

Symmetry is an experimentally verifiable and fundamental property of Quantum Mechanics, see, e.g., the Law of Reciprocity in [4], p. 35. It is satisfied by scalar product of rays. It is also obviously satisfied in classical SP-structures.

Nevertheless Symmetry may be telling us more about our intellectual processes, our logic, than about the structure of the physical world out there. If we accept the idea that states possess the dual aspects of *states the world is in* and of *states we can test for* and that two states are linked by the fundamental $p(x, y)$, rejecting Symmetry would be akin to rejecting the idea that those dual aspects of states are aspects of the same entity, and imply we are dealing with two different types of entities.

4.2 Non-negativity

Our second assumption is that the similarity p is nonnegative:

Property 2 (Non-negativity) *For any $x, y \in \Omega$, $p(x, y) \geq 0$.*

This is requested by the interpretation of p as a “probability” and is obviously satisfied by the scalar product of rays and in classical SP-structures. It seems that Non-negativity does not tell us anything about the physical world but is a logical requirement following from the way our experiments are built.

Since 0 has a special meaning, as the smallest possible value for p , it is natural to pay special attention to those pairs x, y for which $p(x, y) = 0$. Following common usage we shall say that x and y are orthogonal and write $x \perp y$ iff $p(x, y) = 0$. Note that $y \perp x$ iff $x \perp y$, by Symmetry. Similarly we shall say that x is orthogonal to a set A of states and write $x \perp A$ iff $p(x, A) = 0$. Note that for any state x , $x \perp \emptyset$. We shall use the notation $B \perp A$ to mean: for every $x \in A, y \in B$, one has: $x \perp y$.

Definition 2 *A set A of states will be called an ortho-set iff any two distinct elements of A are orthogonal: for any $x, y \in A$ such that $x \neq y$, one has $x \perp y$.*

Note that the empty set is an ortho-set and so is any singleton set. Ortho-sets play a central role in our analysis. They represent states that correspond to different values of an observable physical quantity.

We shall now generalize p to accept not a single state, but any ortho-set of states as a second argument. If $x \in \Omega$ and $A \subseteq \Omega$ is an ortho-set, we define

$$p(x, A) = \sum_{y \in A} p(x, y).$$

The (finite or infinite) sum above is independent of the order of summation. Note that $p(x, A)$ is either a nonnegative real number or $+\infty$.

5 Boundedness, subspaces

We may now introduce our next requirement.

Property 3 (Boundedness) *For any state $x \in \Omega$ and any ortho-set A , $p(x, A) \leq 1$.*

The last inequality should be understood as: $p(x, A)$ is finite and at most one. Again this is a fundamental property in Quantum Physics. The elements of

an ortho-set A represent different possible answers to a unique test. The sum of the “probabilities” of obtaining certain answers cannot be greater than one. Boundedness is satisfied both in Hilbert and in classical SP-structures. Again, Boundedness seems to be a logical requirement, following from our interpretation of orthogonal states as corresponding to different values and of similarity as a transition probability.

Definition 3 *If A is an ortho-set, the subspace $\bar{A} \subseteq \Omega$ generated by A is defined by: $\bar{A} = \{x \in \Omega \mid p(x, A) = 1\}$. The ortho-set A is said to be a basis for \bar{A} . A basis is a basis for Ω . A subspace is a set of states $X \subseteq \Omega$ such that there exists some ortho-set A such that $y = \bar{A}$.*

In classical structures $\bar{A} = A$.

In the following lemma, and throughout this paper we shall assume that the structure $\langle \Omega, p \rangle$ satisfies all the assumptions previously made. In Lemma 1, therefore, p is assumed to satisfy Symmetry, Non-negativity and Boundedness.

Lemma 1 *Let A be an ortho-set. For any $x \in \Omega$, $p(x, A) \in [0, 1]$. In particular, $p(x, y) \in [0, 1]$ for any $y \in \Omega$.*

Proof: By Non-negativity, we have $p(x, A) \geq 0$. By Boundedness, $p(x, A) \leq 1$. The singleton $\{y\}$ is an ortho-set and therefore $p(x, y) = p(x, \{y\}) \in [0, 1]$. ■

Any state orthogonal to each of the states of an ortho-set A is orthogonal to every state in the subspace generated by A .

Lemma 2 *Suppose $x \in \Omega$ is a state and $A \subseteq \Omega$ is an ortho-set such that $x \perp A$. Then, $x \perp \bar{A}$.*

Proof: Since A is an ortho-set and we have $p(x, A) = 0$, the set $A \cup \{x\}$ is an ortho-set. By Boundedness and Symmetry then we have: for any $y \in \bar{A}$ $p(y, A) + p(y, x) \leq 1$. But $p(y, A) = 1$ and therefore $p(y, x) = 0$. ■

6 O-Projection and consequences

6.1 O-projection

The next property we want to consider deals with orthogonal projections.

Property 4 (O-Projection) Suppose $x \in \Omega$ is a state and $A \subseteq \Omega$ is an ortho-set such that $p(x, A) < 1$. Then there exists a state $y \in \Omega$ with the following properties:

1. $y \perp A$, i.e., $p(y, A) = 0$, i.e., $A \cup \{y\}$ is an ortho-set, and
2. $p(x, A) + p(x, y) = 1$.

O-Projection should remind the reader of the Gram-Schmidt process. Physically, the ortho-set A represents certain values of a given observable and therefore can be interpreted as a test: is the state x in A or not. If $p(x, A) < 1$ the answer to the question above may, with a certain “probability” be “no”. If the answer is indeed “no” the system is left in a state y that satisfies the three conditions above. The scalar product can be seen to satisfy those conditions, when y is the projection of x on the subspace A^\perp orthogonal to A . In a classical system, $p(x, A) < 1$ implies $p(x, A) = 0$ and we can take $y = x$. The conditions of O-Projection seem to be logical requirements.

Lemma 3 For any states $x, y \in \Omega$, if $x \sim y$ then $p(x, y) = 1$. In particular, $p(x, x) = 1$.

Proof:

Since $\{y\}$ is an ortho-set, if it were the case that $p(x, y) < 1$, there would exist, by O-Projection, some state z such that $p(y, z) = 0$ and $p(x, y) + p(x, z) = 1$. But $p(x, z) = p(y, z) = 0$ and we conclude that $p(x, y) = 1$.

■

6.2 Bases: existence and size

Lemma 4 Let A be some ortho-set and assume that $B \subseteq \bar{A}$ is such that, for every $x \in \bar{A}$, $p(x, B) = 1$, then B is a basis for \bar{A} .

Proof: We only need to show that B is an ortho-set. Let $x, y \in B$, $x \neq y$. We have $p(x, B) = 1$. But $1 = p(x, B) \geq p(x, x) + p(x, y)$ by Non-negativity. But, by Lemma 3, $p(x, x) = 1$ and we have $p(x, y) = 0$, by Non-negativity.

■

Theorem 1 Let A be an ortho-set. Then there is a basis B such that $A \subseteq B$.

Proof: By ordinal induction, we define an ortho-set set $B_\alpha \subseteq \Omega$ for every ordinal α . We let $B_0 = A$. For a limit ordinal α we set $B_\alpha = \bigcup_{\beta < \alpha} B_\beta$. For any successor ordinal $\alpha + 1$, if B_α is a basis we set $B_{\alpha+1} = B_\alpha$, and if B_α is not a basis, we consider some state $x \in X$ such that $p(x, B_\alpha) < 1$ and we set $B_{\alpha+1} = B_\alpha \cup \{y\}$, where $y \in \Omega$ is one of the states the existence of which is guaranteed by O-Projection. Clearly we have a chain of ortho-sets and there is some ordinal β for which $B_{\beta+1} = B_\beta$. The set B_α is a basis. ■

It is a striking property of Hilbert spaces that any two bases have the same cardinality. The same holds in SP-structures.

Theorem 2 *Let A, B be orthosets such that $B \subseteq \bar{A}$ and assume A is finite. Then B is finite and $|B| \leq |A|$. An SP-structure that admits a finite basis, will be called finite-dimensional and its dimension is the (common) size of its bases.*

Proof: We have $\sum_{b \in B} \sum_{a \in A} p(a, b) = \sum_{b \in B} 1 = |B|$. But $\sum_{a \in A} \sum_{b \in B} p(a, b) \leq \sum_{a \in A} 1 = |A|$. We conclude that $|B| \leq |A|$. ■

We may now define a natural operation on subspaces: orthogonal complement.

Theorem 3 *Let X be any subspace. The set $X^\perp = \{x \in \Omega \mid x \perp X\}$ is a subspace and $X = (X^\perp)^\perp$.*

Proof: Let A be a basis for X . Complete A to a basis $A \cup B$, with B an ortho-set orthogonal to A . We shall show that B is a basis for X^\perp . First, $B \subseteq X^\perp$ by Lemma 2. But, for any state x of X^\perp , $p(x, A) + p(x, B) = 1$ and $p(x, A) = 0$. One sees that A is a basis for $(X^\perp)^\perp$. ■

6.3 Projections on subspaces

The following defines projections on subspaces.

Lemma 5 *If x is a state and A is an ortho-set, such that $p(x, A) > 0$ there is a state y , such that:*

1. $y \in \bar{A}$, and
2. $p(x, y) = p(x, A)$.

Proof: Let B be a basis such that $B = A \cup C$ with $C \perp A$. The existence of such a basis follows from Theorem 1. The set C is an ortho-set and $p(x, A) + p(x, C) = 1$. Since $p(x, A) > 0$, we have $p(x, C) < 1$, and, by O-Projection there is a state y such that $y \perp C$, $p(x, C) + p(x, y) = 1$. But $p(y, A) + p(y, C) = 1$ and $p(y, C) = 0$. Therefore $p(y, A) = 1$. Also $p(x, y) = 1 - p(x, C) p(x, A)$. ■

The reader may wonder about the case $p(x, A) = 0$. In this case, by Lemma 2, every state y such that $p(y, A) = 1$ satisfies the condition required, i.e., $p(x, y) = 0$.

7 Factorization and Consequences

Our next defining property for SP-structure is a factorization property.

Property 5 (Factorization) *Let A be an ortho-set and x an arbitrary state. If $y, z \in \bar{A}$ and $p(x, y) = p(x, A)$, then $p(x, z) = p(x, y) p(y, z)$.*

Factorization implies that $p(x, A)$ is the maximum of all $p(x, y)$ for $y \in \bar{A}$ and that every such $p(x, y)$ can be factored out through the state taking this maximum. Factorization has been described in Theorem 1 of [2]. The meaning of Factorization, for Physics, is that, if one knows that in state y some observable A has a specific value, then the probability of a transition from x to y is the product of the probability of measuring this specific value (in x) times the transition probability from the state obtained after the measurement to y . Factorization seems to be a logical requirement relating tests to two propositions one of which entails the other: if A entails B , testing for A may be done by testing first for B and then for A .

Theorem 4 *For any state $x \in \Omega$ and any ortho-set A , $p(x, A) = \max(\{p(x, y) \mid y \in \bar{A}\})$. Therefore if B is an ortho-set such that $\bar{B} = \bar{A}$, one has $p(x, A) = p(x, B)$. From now on, if X is a subspace we shall allow ourselves the use of the notation $p(x, X)$. Also, if X and Y are subspaces such that $X \subseteq Y$, then, for any $x \in \Omega$, one has $p(x, X) = p(x, Y) p(t(x, Y), X)$.*

In the last equation note that in the case $t(x, Y)$ is not defined, we have $p(x, Y) = 0$ and therefore we consider the product on the right hand side of the last equation to be defined and equal to zero.

Proof: By Lemma 5 there is some $z \in \bar{A}$ such that $p(x, A) = p(x, z)$ and by Factorization we have, for every $y \in \bar{A}$, $p(x, y) = p(x, z) p(z, y) \leq p(x, z)$. The remainder follows easily. ■

The Factorization property has many consequences that will be presented now. The first one concerns the relation of equivalence between states.

7.1 Similarity and Equivalence

Lemma 6 *If $a \perp b$, $x \perp a$ and $p(y, a) + p(y, b) = 1$, then we have $p(x, y) = p(x, b) p(y, b)$.*

Proof: Let $A = \{a, b\}$. The set A is an ortho-set. If $p(x, b) = 0$, $p(x, A) = 0$ and, by Lemma 2, $p(x, y) = 0$ and the claim is proved.

Assume, then, that $p(x, b) > 0$. Note that $p(b, A) = 1$ and $p(x, A) = p(x, b)$. By Factorization, then $p(x, y) = p(x, b) p(b, y)$. ■

Theorem 5 *Any states $x, y \in \Omega$ are equivalent, i.e., $x \sim y$, iff $p(x, y) = 1$.*

Proof: If $x \sim y$, $p(x, y) = p(x, x)$, and we conclude the proof with Lemma 3.

Suppose, now, that $p(x, y) = 1$ and that $z \in \Omega$. We want to show that $p(x, z) = p(y, z)$. Without loss of generality, we can assume $p(z, x) < 1$. By O-Projection, there is some state $x' \perp x$ such that $p(z, x) + p(z, x') = 1$. By Boundedness we have $p(y, x) + p(y, x') \leq 1$ and therefore $p(y, x') = 0$. The assumptions of Lemma 6 are satisfied for $a = x'$, $b = x$, $x = y$ and $y = z$. We conclude that $p(y, z) = p(y, x) p(z, x) = p(z, x)$. ■

Theorem 5 shows that, if $p(x, y) = 1$, then x and y are equivalent, i.e., behave in exactly the same way as far as p is concerned. No harm can therefore be caused by identifying any two states x, y such that $p(x, y) = 1$.

Definition 4 *An SP-structure $\langle \Omega, p \rangle$ is said to be standard iff for any $x, y \in \Omega$, $p(x, y) = 1$ implies $x = y$.*

Theorem 6 *Let $\langle \Omega, p \rangle$ be an SP-structure. The quotient structure $\langle \Omega / \sim, \bar{p} \rangle$ defined by $\bar{p}(\bar{x}, \bar{y}) = p(x, y)$ is a standard SP-structure and the transformation $x \mapsto \bar{x}$ preserves p .*

In the sequel we shall only consider standard SP-structures, even if we forget to mention the fact. In other words, we assume, from now on, that $p(x, y) = 1$ iff $x = y$.

7.2 Relativization

We shall also strengthen O-Projection and Theorem 1 by relativizing them to a subspace. The relativization of O-Projection shows that any subspace of an SP-structure is an SP-structure. First, we need the following.

Lemma 7 *Let x, x' be states such that $p(x, x') < 1$. Let y be a state orthogonal to x' such that $p(x, x') + p(x, y) = 1$ as guaranteed by O-Projection. Then, for any ortho-set A such that $x \perp A$ and $x' \perp A$, we have $y \perp A$.*

Proof: Let z be a state in A . Since $z \perp x'$, by factorization we have $p(x, z) = p(x, y)p(y, z)$. But $x \perp z$ and $p(x, z) = 0$. But $p(x, y) > 0$ and therefore $p(y, z) = 0$. ■

Theorem 7 *Suppose $X \subseteq \Omega$ is a subspace, $x \in X$ is a state and $A \subseteq X$ is an ortho-set of X such that $p(x, A) < 1$. Then any state $y \in \Omega$ such that*

1. $y \perp A$, i.e., $p(y, A) = 0$, i.e., $A \cup \{y\}$ is an ortho-set, and
2. $p(x, A) + p(x, y) = 1$

is a member of X .

Proof: We have $A \perp X^\perp$ and $x \perp X^\perp$ and therefore, by Lemma 7, $y \perp X^\perp$ and, by Theorem 3, $y \in X$. ■

We may also relativize Theorem 1

Theorem 8 *Let A be an ortho-set and X be a subspace such that $A \subseteq X$. Then there is a basis B for X such that $A \subseteq B$.*

The proof follows that of Theorem 1, using Theorem 7 to show that $B_\alpha \subseteq X$.

7.3 Intersections and orthogonal sums of subspaces

In Theorem 3 we defined orthogonal complements for subspaces. Once we have established the meaning of $p(x, X)$ for a subspace X , as has been done in Theorem 4, the following is an obvious corollary of Theorem 3.

Corollary 1 *For any $x \in \Omega$ and any subspace X , one has $p(x, X) + p(x, X^\perp) = 1$.*

The proof is obvious.

We may now define other natural operations on subspaces. First we define orthogonal sums.

Theorem 9 *Let $X \perp Y$ be orthogonal subspaces. The set $X \oplus Y$, defined to be $\{x \in \Omega \mid p(x, X) + p(x, Y) = 1\}$, is a subspace.*

Proof: Let A, B be bases for X, Y respectively. The ortho-set $A \cup B$ is a basis for $X \oplus Y$ by Theorem 4. ■

We may now define intersections.

Theorem 10 *If X and Y are subspaces their intersection $X \cap Y$ is also a subspace.*

Proof: We shall build an ortho-set $A_\alpha \subseteq X \cap Y$ for every ordinal α , by ordinal induction. Let $A_0 = \emptyset$. For a limit ordinal α we set $A_\alpha = \bigcup_{\beta < \alpha} A_\beta$. For any successor ordinal $\alpha + 1$, if A_α is a basis for $X \cap Y$, we set $A_{\alpha+1} = A_\alpha$, and if A_α is not a basis for $X \cap Y$ we consider some state $x \in X \cap Y$ such that $p(x, A_\alpha) < 1$ and we set $A_{\alpha+1} = A_\alpha \cup \{y\}$, where $y \in \Omega$ is one of the states the existence of which is guaranteed by O-Projection. By Theorem 7, $y \in X \cap Y$. Clearly we have a chain of ortho-sets and there is some ordinal β for which $A_{\beta+1} = A_\beta$. The set A_α is a basis for $X \cap Y$. ■

7.4 Uniqueness of projections

We may now strengthen O-Projection and Lemma 5. We then prove a fundamental result on subspaces: the projections guaranteed by O-Projection and Lemma 5 are unique and independent of the basis considered.

Theorem 11 *In a standard SP-structure, if x is a state and A is an ortho-set, such that $p(x, A) > 0$, the state $y \in \bar{A}$ such that $p(x, y) = p(x, A)$ guaranteed by Lemma 5 is unique and depends only on \bar{A} , not on A . This unique state will be denoted $t(x, \bar{A})$ or by $t(x, A)$.*

Note that $t(x, A)$ is defined only if $p(x, A) > 0$ and that: $p(t(x, A), A) = 1$, $p(x, t(x, A)) = p(x, A)$ and for any $y \in \Omega$ such that $p(y, A) = 1$ one has $p(x, y) = p(x, t(x, A)) p(t(x, A), y)$.

Proof: Suppose both y_i $i = 0, 1$ satisfy the conditions. By Factorization we have $p(x, y_i) = p(x, y_{i+1}) p(y_{i+1}, y_i)$ for $i = 0, 1$ where $1 + 1 = 0$. We conclude

that $p(y_0, y_1) = 1$ and, by Theorem 5, that $y_0 \sim y_1$. Suppose now that B is an ortho-set such that $\bar{B} = \bar{A}$. We have $y_0 \in \bar{B}$ and $p(x, y_0) = p(x, B)$ by Theorem 4. Therefore y_0 is the projection of x on B . ■

7.5 Completion is a closure operation

The following shows that the completion of an ortho-set into the subspace it generates has the character of a closure operation.

Theorem 12 *Let $A, B \subseteq \Omega$ be ortho-sets. The following properties are equivalent:*

1. $\bar{A} \subseteq \bar{B}$,
2. $A \subseteq \bar{B}$,
3. for any $x \in \Omega$ $p(x, A) \leq p(x, B)$.

Proof: Item 1 clearly implies item 2 since $A \subseteq \bar{A}$. Item 3 implies item 1 by Boundedness: $1 = p(x, A) \leq p(x, B) \leq 1$. Let us show that item 2 implies item 3. Assume $A \subseteq \bar{B}$ and $x \in \Omega$. The ortho-set A can be extended into a basis for the subspace \bar{B} by Theorem 1. Therefore $p(x, A) \leq p(x, \bar{B}) = p(x, B)$ by Theorem 11. ■

7.6 An iterative description of O-Projection

We can also strengthen O-Projection.

Theorem 13 *If A is an ortho-set and x a state such that $p(x, A) < 1$ then there is a unique state y such that $y \perp A$ and $p(x, A) + p(x, y) = 1$. This state y is $t(x, \bar{A}^\perp)$ and therefore depends only on \bar{A} and not on A .*

Proof: Suppose $y_i \perp A$ and $p(x, A) + p(x, y_i) = 1$ for $i = 0, 1$. We have $y_i \in A^\perp$ and $p(x, y_i) = p(x, A^\perp)$ by Lemma 1. We conclude by Theorem 11. ■

Property 4, O-Projection claims the existence of the projection of a state x on the subspace A^\perp orthogonal to any ortho-set (in fact any subspace) A . Could we have weakened our assumption and required only the existence of such a projection when the ortho-set A is a single state? The answer is

negative: for infinite ortho-sets A , i.e., for infinite-dimensional subspaces the full force of O-Projection is needed. But we shall show now that, for finite-dimensional subspaces, the existence of o-projections follows from the simple case of a one-dimensional space, with the help of Property 5, Factorization.

Theorem 14 *Let $A = \{a\} \cup A'$ be an ortho-set. Assume $p(x, A) < 1$. Then, $n(x, A) = n(n(x, A'), \{a\})$.*

Proof: Assume $p(x, A) < 1$. By Non-negativity, $p(x, A') < 1$ and, by O-Projection on A' the state $n(x, A')$ exists and we have $n(x, A') \perp A'$, $p(x, A') + p(x, n(x, A')) = 1$ and, by Factorization, for any state $y \perp A'$ one has $p(x, y) = p(x, n(x, A')) p(n(x, A'), y)$. We conclude, first, that $p(x, a) < p(x, n(x, A'))$ and, by Theorem 5, $p(a, n(x, A')) < 1$. Similarly, we see that $p(a, n(x, A')) < 1$. Therefore, $b = n(n(x, A'), a)$ is well-defined and we have $b \perp a$, $p(n(x, A'), a) + p(n(x, A'), b) = 1$ and for any state $y \perp a$ one has $p(n(x, A'), y) = p(n(x, A'), b) p(b, y)$. We notice, first, that $p(n(x, A'), b) > 0$ since $p(n(x, A'), a) < 1$. For any $w \in A'$ we have $p(n(x, A'), w) = p(n(x, A'), b) p(b, w)$. But $n(x, A') \perp w$ and $p(n(x, A'), b) > 0$ and therefore $b \perp w$. We have shown that $b \perp A$.

We shall now prove that $p(x, A) + p(x, b) = 1$. We have:

$$\begin{aligned} p(x, A) + p(x, b) &= p(x, a) + p(x, A') + p(x, b) = \\ p(x, n(x, A')) p(n(x, A'), a) + p(x, A') + p(x, n(x, A')) p(n(x, A'), b) &= \\ p(x, n(x, A')) (p(n(x, A'), a) + p(n(x, A'), b)) + p(x, A') &= \\ p(x, n(x, A')) + p(x, A') &= 1 \end{aligned}$$

■

When the ortho-set A of O-Projection is finite, then the state y can be described by a sequence of O-projections in which the ortho-set considered is a singleton.

8 Notations

Before we can express our next requirement we need some notations. We wish to consider the following general situation. Suppose $X \perp Y$ are orthogonal subspaces and let $Z = X \oplus Y$ be their orthogonal sum. Assume a and b

are states that are not orthogonal to Z , i.e., equivalently, none of a or b is orthogonal to both X and Y . The states a and b have projections on Z : $t(a, Z)$ and $t(b, Z)$.

Theorem 15 *Let $X \perp Y$ be subspaces and $Z = X \oplus Y$. Assume $p(a, Z) > 0$. If $p(a, X) > 0$, a has a projection on X and $t(a, X) = t(t(a, Z), X)$.*

Proof: By Theorem 11 since $p(a, t(a, X)) = p(a, t(a, Z) p(t(a, Z), t(a, X)))$
■

Suppose $X \perp Y$, $Z = X \oplus Y$ and $a, b \in Z$. We expect the quantity $p(a, b)$ to be related to $p(t(a, X), t(b, X))$ and $p(t(a, Y), t(b, Y))$. We shall therefore define two related quantities.

Definition 5 *Let $X \perp Y$, $Z = X \oplus Y$, $a, b \in \Omega$, $p(a, Z) > 0$, $p(b, Z) > 0$, we shall define:*

$$\begin{aligned} (1) \quad \alpha_{X,Y}(a, b) &= p(t(a, Z), t(b, Z)) - \\ &\quad p(a, X) p(b, X) p(t(a, X), t(b, X)) - p(a, Y) p(b, Y) p(t(a, Y), t(b, Y)) \\ (2) \quad \rho_{X,Y}(a, b) &= \\ &\quad 2 \sqrt{p(a, X) p(b, X) p(a, Y) p(b, Y) p(t(a, X), t(b, X)) p(t(a, Y), t(b, Y))} \end{aligned}$$

We use the convention that the product by zero of an undefined quantity is equal to zero. Note, then, that the definitions of α and ρ above are legal even if one or more of the expressions $t(a, X)$, $t(a, Y)$, $t(b, X)$, $t(b, Y)$ is not defined, or, equivalently if a or b is in X or Y (we assumed a and b are not perpendicular to Z). For example, if $t(a, X)$ is not defined, i.e., if $a \in Y$, then $p(a, X) = 0$.

Note that $\rho_{X,Y}(a, b) = 0$ iff a or b is in X or Y . Appendix B describes those quantities in Hilbert spaces.

Lemma 8 *For any $a \in Z = X \oplus Y$ one has $\alpha_{X,Y}(a, a) = \rho_{X,Y}(a, a)$.*

Proof:

For any $a \in \Omega$ $\rho_{X,Y}(a, a) = 2 p(a, X) p(a, Y)$ and $\alpha_{X,Y}(a, a) = 1 - p^2(a, X) - p^2(a, Y)$. If $a \in Z$ we have $p(a, X) + p(a, Y) = 1$ and therefore $1 - p^2(a, X) - p^2(a, Y) = 2 p(a, X) p(a, Y)$.
■

Lemma 9 *Let $x \perp y$ be orthogonal states and let X be the subspace generated by x and y . If $a, b \in X$ and $a \perp b$, then $\alpha_{x,y}(a, b) = -\rho_{x,y}(a, b)$.*

Proof: We have $p(a, x) + p(a, y) = 1 = p(b, x) + p(b, y)$. By Boundedness, we have $p(x, a) + p(x, b) \leq 1$ and also $p(y, a) + p(y, b) \leq 1$. We conclude that $p(x, a) + p(x, b) = 1 = p(y, a) + p(y, b)$ and therefore $p(b, y) = p(a, x)$ and $p(a, y) = p(b, x)$. We see that $\rho_{x,y}(a, b) = 2p(a, x)p(b, x)$ and that $\alpha_{x,y}(a, b) = 0 - 2p(a, x)p(b, x) = -\rho_{x,y}(a, b)$. ■

Definition 6 *Let $X \perp Y$ and $a, b \in \Omega$, such that a (resp. b) is orthogonal to neither X nor Y . The projections $t(a, X)$ and $t(a, Y)$ (resp. $t(b, X)$ and $t(b, Y)$) are well-defined and $\rho_{X,Y}(a, b) > 0$. We shall define*

$$\omega_{X,Y}(a, b) = \frac{\alpha_{X,Y}(a, b)}{\rho_{X,Y}(a, b)}$$

9 A fundamental inequality

We present our next assumption.

Property 6 (Inequality) *For any orthogonal subspaces $X \perp Y$ and any states $x, y \in Z = X \oplus Y$, one has $|\alpha_{X,Y}(a, b)| \leq \rho_{X,Y}(a, b)$. In other words, if $\rho_{X,Y}(a, b) > 0$, then $-1 \leq \omega_{X,Y}(a, b) \leq 1$ and if $\rho_{X,Y}(a, b) = 0$, then $\alpha_{X,Y}(a, b) = 0$.*

The physical meaning of the property above is not fully understood at this point. It must be related to the two-paths experiments that are so central in Quantum Physics. The deep meaning of Inequality is probably hidden in Theorem 17 that shows that it implies a sort of continuity property: if $p(x, y)$ is close to 1, then x and y are almost equivalent.

In classical SP-structures, Appendix A shows that for any X, Y, a, b we have $\alpha_{X,Y}(a, b) = 0 = \rho_{X,Y}(a, b)$. Appendix B studies the Hilbert SP-structures. In an SP-structure defined by a Hilbert space on the real field, $\rho_{X,Y}(a, b)$ can be any number in the interval $[0, 1/2]$ and $\alpha_{X,Y}(a, b)$ is either equal to $\rho_{X,Y}(a, b)$ or equal to $-\rho_{X,Y}(a, b)$. But we can say more. Let us say that a and b are parallel if $\alpha_{X,Y}(a, b) = \rho_{X,Y}(a, b) > 0$ and say that they are opposite if $-\alpha_{X,Y}(a, b) = \rho_{X,Y}(a, b) > 0$. Assume now that a and c are parallel. If c and b are parallel, then a and b are parallel and if c and b are opposite then a and b are opposite.

In an SP-structure defined by a Hilbert space on the complex field, one can define an angle $\varphi_{X,Y}(a, b) \in [0, 2\pi[$ such that, for any states a, b and c that are neither in X nor in Y one has:

1. $\varphi_{X,Y}(b, a) = -\varphi_{X,Y}(a, b),$
2. $\varphi_{X,Y}(a, b) = \varphi_{X,Y}(a, c) + \varphi_{X,Y}(c, b),$
3. $\alpha_{X,Y}(a, b) = \rho_{X,Y}(a, b) \cos(\varphi_{X,Y}(a, b)).$

All angles may appear and a state a is uniquely characterized by $t(a, X)$, $t(a, Y)$, $p(a, X)$ and an arbitrary angle $\psi(a)$ defined up to an additive constant. Then, $\varphi_{X,Y}(a, b) = \psi(b) - \psi(a)$ (or $\psi(a) - \psi(b)$).

In an SP-structure defined by a Hilbert space on the quaternions phases are not angles but unit quaternions.

We see that in each of the four kinds of SP-structures considered above some additional properties hold, but they are different and their physical meaning is unclear. Since our purpose is to define SP-structures in such a way to cover classical and all three kinds of Hilbert space structures, we do not impose any further requirements on SP-structures.

10 Consequences of Inequality

We shall now examine the consequences of our new assumption. First we show that Inequality can be strengthened: the condition $x \in Z$ is superfluous.

Theorem 16 *For any states orthogonal subspaces $X \perp Y$ and states a, b such that $p(b, X) + p(b, Y) = 1$, one has $|\alpha_{X,Y}(a, b)| \leq \rho_{X,Y}(a, b)$.*

Proof: Let $Z = X \oplus Y$ be the orthogonal sum of X and Y . The set Z is a subspace. If $p(a, Z) > 0$, by Theorem 11, we have, for $w = t(a, Z) \in Z$, $p(a, b) = p(a, w)p(w, b)$, $p(a, X) = p(a, w)p(w, X)$, and $p(a, Y) = p(a, w)p(w, Y)$. Therefore $\alpha_{X,Y}(a, b) = p(a, w)\alpha_{X,Y}(w, b)$. Similarly, $\rho_{X,Y}(a, b) = p(a, w)\rho_{X,Y}(w, b)$. By Inequality we have $|\alpha_{X,Y}(w, b)| \leq \rho_{X,Y}(w, b)$. Since $p(a, w) \leq 1$ we conclude that $|\alpha_{X,Y}(a, b)| \leq \rho_{X,Y}(a, b)$.

If $p(a, Z) = 0$, then, by Lemma 2, $p(a, b) = 0$ and $\alpha_{X,Y}(a, b) = 0$. ■

Our next result is a continuity property: if $p(x, y)$ is close to one, then, for any z , $p(x, z)$ is close to $p(y, z)$.

Theorem 17 *For any $x, y, z \in \Omega$, one has:*

$$p(x, z) \leq p(y, z) + 1/2 \sqrt{1 - p(x, y)} + (1 - p(x, y)).$$

Proof: Assume, for now, that there is some state $y' \in \Omega$ such that $y' \perp y$ and $p(x, y) + p(x, y') = 1$. Consider any $z \in \Omega$. By Theorem 16: $|\alpha_{y, y'}(z, x)| \leq \rho_{y, y'}(z, x)$. But $2\alpha_{y, y'}(z, x) = p(z, x) - p(z, y)p(x, y) - p(z, y')p(x, y') \geq p(z, x) - p(z, y) - p(x, y')$. Also, $\rho_{y, y'}(z, x) \leq \sqrt{p(x, y')}$. We conclude that $p(z, x) - p(z, y) \leq 1/2 \sqrt{1 - p(x, y)} + 1 - p(x, y)$.

Now, if $p(x, y) < 1$, there is such a state y' by O-Projection. Assume, then, that $p(x, y) = 1$. If there is some state $y' \in \Omega$ orthogonal to y , we have $1 \geq p(x, y) + p(x, y') = 1 + p(x, y')$, by Boundedness, and we conclude that $p(x, y) + p(x, y') = 1$.

Let us deal now with the limit case: there is no $y' \in \Omega$ that is orthogonal to y . By O-Projection this implies that for every $z \in \Omega$, $p(z, y) = 1$, and our claim is proved. ■

Appendix C shows that the bounds of Theorem 17 are tight.

11 Similarity-preserving mappings

This section presents preliminary results on mappings that preserve similarity. We want, now, to consider morphisms between SP-structures.

Definition 7 *A morphism from SP-structure $S_1 = \langle \Omega_1, p_1 \rangle$ to SP-structure $S_2 = \langle \Omega_2, p_2 \rangle$ is a function $f : \Omega_1 \longrightarrow \Omega_2$ that preserves similarity, i.e., such that: for any $a, b \in \Omega_1$ we have $p_2(f(a), f(b)) = p_1(a, b)$.*

Theorem 18 *Let S_1 be a standard SP-structure. Any morphism f from S_1 to any SP-structure S_2 is injective. Any such morphism that is surjective is an isomorphism: it has an inverse that is a morphism.*

Proof: If $f(a) = f(b)$ then $p_2(f(a), f(b)) = 1$ and $p_1(a, b) = 1$. Since S_1 is standard, we have $a = b$. If f is surjective it is bijective and therefore has an inverse f^{-1} . But $p_1(f^{-1}(a), f^{-1}(b)) = p_2(f(f^{-1}(a)), f(f^{-1}(b))) = p_2(a, b)$. ■

Theorem 19 *If $f : S_1 \longrightarrow S_2$ is an isomorphism, then the direct image by f of any basis of S_1 is a basis of S_2 and the direct image by f of any subspace of S_1 is a subspace of S_2 .*

The proof is obvious.

We are now interested in studying particular isomorphisms. First those isomorphisms are automorphisms, i.e., isomorphisms from an SP structure to itself. Secondly there is a basis of invariant states.

Theorem 20 *Let $S = \langle \Omega, p \rangle$ be a standard SP-structure and let B be a basis for S . Let now $f : S \rightarrow S$ be an isomorphism such that for every element b of the base B , we have $f(b) = b$. Let $A \subseteq B$ and let X be the subspace generated by A . We have:*

1. *for any $x \in \Omega$, $p(x, X) = p(f(x), X)$,*
2. *X is (globally) invariant under f : $f(X) = X$,*
3. *if x is not orthogonal to X then $t(f(x), X) = f(t(x, X))$.*

Proof: We have $p(x, X) = \sum_{a \in A} p(x, a) = \sum_{a \in A} p(f(x), f(a)) = \sum_{a \in A} p(f(x), a) = p(f(x), X)$. We have proved item 1. For item 2, note that, by Theorem 19, $f(X)$ is a subspace. By 1 above, A is a basis for $f(X)$. Therefore $f(X) = \bar{A} = X$. For item 3, assume $p(x, X) > 0$. Then, by 1 above, $p(f(x), X) > 0$ and both $t(x, X)$ and $t(f(x), X)$ are well-defined. By Theorem 11 $t(f(x), X)$ is the unique state in X such that $p(f(x), t(f(x), X)) = p(f(x), X)$. But $p(f(x), X) = p(x, X) = p(x, t(x, X)) = p(f(x), f(t(x, X)))$. ■

12 Observables

12.1 Definition

We now want to study physical properties. In Classical Physics, those are numerical values attached to each possible state: a physical property is represented by a function from the set of all possible states to the real numbers and essentially (except for some continuity condition in the infinite case) any such function represents a possible physical property. In Quantum Physics, such physical properties, often called *observables*, are represented by Hermitian, i.e., self-adjoint, operators. Such operators have, for Quantum Physics, a triple role, that we want to analyze here.

- First, as in the classical case, they attach, through their eigenvalues, values to states. Those values, in the Quantum case, are interpreted as mean values.

- But, secondly, those operators define eigensubspaces and projections on those that characterize the change of state caused by a measurement,
- and, thirdly, they represent linear transformations in the state space, typically interpreted as infinitesimal transformations whose commutation properties are significant.

In many respects such physical properties (or observables) behave in a way that resembles random variables and therefore maybe they should be termed *variables* instead of observables.

Since the third aspect above is central in Quantum Physics, we shall present observables as transformations of a special kind on SP-structures. Observables, viewed as transformations, correspond to Hermitian, i.e., self-adjoint bounded linear operators.

Definition 8 *Let $\langle \Omega, p \rangle$ be an SP-structure. An observable (of the SP-structure) is a transformation $r : \Omega \longrightarrow \Omega$ satisfying the following conditions:*

1. *there exists a decomposition of Ω in a denumerable set of non-empty pairwise orthogonal subspaces, i.e., there is a denumerable set I and non-empty subspaces X_j , for every $j \in I$ such that:*
 - *for any $j, k \in I$, if $j \neq k$ then $X_j \perp X_k$ and*
 - $\Omega = X_1 \oplus \dots \oplus X_j \oplus \dots$,
2. *for every $j \in I$, there is a real number λ_j such that*
 - *the λ 's are pairwise different, i.e., for any $j, k \in I$, if $j \neq k$ then $\lambda_j \neq \lambda_k$,*
 - *the λ 's are bounded, i.e., there some real number M such that for any j , $1 \leq j \leq n$, we have $|\lambda_j| \leq M$,*
3. *for any $a \in \Omega$, for any i , $1 \leq i \leq n$, and for any $b \in X_i$, we have, if $\sum_{j \in I} \lambda_j^2 p(a, X_j) > 0$:*

$$p(r(a), b) = \frac{\lambda_i^2 p(a, b)}{\sum_{j \in I} \lambda_j^2 p(a, X_j)},$$

and if $\sum_{j \in I} \lambda_j^2 p(a, X_j) = 0$ we have $p(r(a), b) = p(a, b)$,

4. for any $j, k \in I$ such that $j \neq k$ and for any $a, b \in \Omega$ such that $\omega_{X_j, X_k}(a, b)$ is defined

(a) if $\lambda_j \lambda_k > 0$, then

$$\omega_{X_j, X_k}(r(a), b) = \omega_{X_j, X_k}(a, b)$$

and,

(b) if $\lambda_j \lambda_k < 0$, then

$$\omega_{X_j, X_k}(r(a), b) = -\omega_{X_j, X_k}(a, b).$$

In connection with item 3, note that $\sum_{j=1}^n \lambda_j^2 p(a, X_j) = 0$ implies that for any $1 \leq i \leq n$ we have $\lambda_i^2 p(a, X_i) = 0$. The λ 's are called eigenvalues. The subspace X_j is the eigensubspace corresponding to the eigenvalue λ_i (remember: eigenvalues are pairwise different). The assumption that eigenvalues are bounded ensures the convergence of the denominator in condition 3. Note that, if r is an observable with eigenvalues λ_i , it is also an observable with eigenvalues $c \lambda_i$ for any $c \neq 0$. Eigenvalues are defined only up to a (non-zero) multiplicative constant. Appendix D shows that Hermitian bounded operators in Hilbert spaces define *observables*.

In a classical SP-structure, the orthogonal sum is set union and any state is an eigenvector. An observable is an arbitrary real bounded function on states, defined up to a multiplicative non-zero constant.

12.2 Eigensubspaces

As expected eigensubspaces corresponding to different eigenvalues are orthogonal. Projections on eigensubspaces are defined as in Theorem 11.

Theorem 21 *Assume r is an observable with eigensubspaces X_j and eigenvalues λ_j for $j \in I$.*

1. *For any $x \in \Omega$, one has $\sum_{j \in I} \lambda_j^2 p(x, X_j) = 0$ iff there exists some $i \in I$ such that $x \in X_i$ and $\lambda_i = 0$.*
2. *For any $x \in \Omega$ and any $i \in I$, if $x \perp X_i$, then $r(x) \perp X_i$.*
3. *For any $x \in \Omega$ and any $i \in I$ one has $r(x) \perp X_i$ iff $x \perp X_i$ or $\lambda_i = 0$ and $x \notin X_i$.*

4. For any $x \in \Omega$ and any $i \in I$, if $r(x) \not\perp X_i$, then both $t(x, X_i)$ and $t(r(x), X_i)$ are defined and they are equal.
5. For any $i \in I$ and any $x \in X_i$, we have $r(x) = x$.
6. For any $x \in \Omega$ and any $i \in I$, if $\sum_{j \in I} \lambda_j^2 p(x, X_j) > 0$, then

$$p(r(x), X_i) = \frac{\lambda_i^2 p(x, X_i)}{\sum_{j \in I} \lambda_j^2 p(x, X_j)}$$

and if $\sum_{j \in I} \lambda_j^2 p(x, X_j) = 0$, then $p(r(x), X_i) = p(x, X_i)$.

7. For any $x \in \Omega$, $r(x) = x$ iff x is an eigenvector, i.e., a member of some X_j .

Proof: Let r be an observable.

For item 1, the *if* part is obvious. For the *only if* part, note that, if $\sum_{j \in I} \lambda_j^2 p(x, X_j) = 0$, we must have, for every $j \in I$, $\lambda_j^2 p(x, X_j) = 0$. Therefore $p(x, X_i) > 0$ implies $\lambda_i = 0$. Since all λ 's are different there is at most one λ_i equal to zero and there is some $i \in I$ such that $\lambda_i = 0$ and $p(x, X_i) = 1$.

In the remainder of this proof we shall use 3 of Definition 8 often and without always mentioning it.

For item 2, note that, if $p(x, a) = 0$ for $a \in X_i$, then, by 3 of Definition 8, $p(r(x), a) = 0$. Therefore $x \perp X_i$ implies $r(x) \perp X_i$.

Consider item 3, now. For the *if* part, use item 2 above and notice that if $p(x, X_i) < 1$ and $\lambda_i = 0$ then $\gamma = \sum_{j \in I} \lambda_j^2 p(x, X_j) > 0$ and therefore, for any $a \in X_i$ we have $p(r(x), a) = (0 p(x, a) / \gamma) = 0$. For the *only if* part assume $p(r(x), X_i) = 0$ and $p(x, X_i) > 0$. There is some state $a \in X_i$ such that $p(x, a) > 0$. But $p(r(x), a) = 0$ and therefore $\lambda_i = 0$ and $\sum_{j \in I} \lambda_j^2 p(x, X_j) > 0$, which implies $x \notin X_i$.

For item 4, assume $p(r(x), X_i) > 0$. By item 2 above, $p(x, X_i) > 0$. Therefore both $t(r(x), X_i)$ and $t(x, X_i)$ are well-defined. The state $t(r(x), X_i)$ is the only state of X_i such that $p(r(x), t(r(x), X_i)) \leq p(r(x), a)$ for every state $a \in X_i$. But $p(x, t(x, X_i)) \leq p(x, a)$ for every such state and

$$p(r(x), t(x, X_i)) = \frac{\lambda_i p(x, t(x, X_i))}{\sum_{j \in I} \lambda_j^2 p(a, X_j)} \geq \frac{\lambda_i p(x, t(x, a))}{\sum_{j \in I} \lambda_j^2 p(a, X_j)} = p(r(x), t(x, a))$$

for any $a \in X_i$. We conclude that $t(r(x), X_i) = t(x, X_i)$.

For item 5, suppose that x is an element of X_i . The state x is orthogonal to any X_j for $j \neq i$. Therefore $r(x)$ is orthogonal to any such X_j by item 2. We see that $r(x) \in X_i$. By item 4 we have

$$r(x) = t(r(x), X_i) = t(x, X_i) = x.$$

Consider item 6, now. The *if* part is item 5 above. For the *only if* part, let us assume that $r(x) = x$. By contradiction, assume that there are $i, j \in I$, $i \neq j$ such that $x \notin X_i$ and $x \notin X_j$. By item 1, $\sum_{j \in I} \lambda_j^2 p(x, X_j) > 0$. By item 6, $p(r(x), X_i) = \frac{\lambda_i^2 p(x, X_i)}{\sum_{j \in I} \lambda_j^2 p(x, X_j)}$ and $p(r(x), X_j) = \frac{\lambda_j^2 p(x, X_j)}{\sum_{j \in I} \lambda_j^2 p(x, X_j)}$. But $r(x) = x$, $p(x, X_i) > 0$ and $p(x, X_j) > 0$ therefore $\lambda_i^2 = \lambda_j^2 = \sum_{j \in I} \lambda_j^2 p(x, X_j)$. Since the λ 's are different, we conclude that $\lambda_i = -\lambda_j$. We then see that there can only be one pair of indexes i, j such that $x \notin X_i$ and $x \notin X_j$. We conclude that $x \in Z = X_i \oplus X_j$. Since $\lambda_i \lambda_j < 0$, by 4 of Definition 8, we have $\omega_{X_i, X_j}(x, a) = -\omega_{X_i, X_j}(x, a)$ for any $a \in \Omega$ and therefore $\omega_{X_i, X_j}(x, a) = 0$ for any state a . But, by Lemma 8 $\omega_{X_i, X_j}(x, x) = 1$. ■

12.3 Mean values

The first role of observables: defining mean values, will be discussed now.

Definition 9 Let r be an observable of the SP-structure $\langle \Omega, p \rangle$ and let $x \in \Omega$ be a state. The mean value of r in (or at) x , denoted $\hat{r}(x)$ is defined by: $\hat{r}(x) = \sum_{i \in I} \lambda_i p(x, X_i)$. This, possibly infinite, sum is absolutely convergent, since the λ 's are bounded.

The definition of the mean value resembles an expected value if one interprets p as a conditional probability.

Theorem 22 Assume r is an observable and B_i , $i \in I$ is a basis consisting of eigenvectors, then for any state $x \in \Omega$, we have $\hat{r}(x) = \sum_{i \in I} \hat{r}(b_i) p(x, b_i)$.

Proof: The states b_i that are elements of X_j form a basis for X_j . By Factorization, then,

$$\begin{aligned} \sum_{i, b_i \in X_j} \hat{r}(b_i) p(x, b_i) &= \sum_{i, b_i \in X_j} \lambda_j p(x, b_i) = \lambda_j \sum_{i, b_i \in X_j} p(x, X_j) p(t(x, X_j), b_i) = \\ &\lambda_j p(x, X_j) \sum_{i, b_i \in X_j} p(t(x, X_j), b_i) = \lambda_j p(x, X_j) \end{aligned}$$

■

Note that Theorem 22 assumes the basis chosen includes only eigenvectors of the observable r . This condition cannot be dispensed with.

The following shows that mean-value functions are, in a sense, continuous. Our claim deals only with finite-dimensional structures.

Theorem 23 *Let Ω be a finite-dimensional SP-structure. Let r be an observable. For any $x, y \in \Omega$,*

$$\hat{r}(x) - \hat{r}(y) \leq \left(\sum_{i \in I} \lambda_i \right) (1/2 \sqrt{1 - p(x, y)} + (1 - p(x, y)))$$

The sum of the eigenvalues is well defined since I is finite, by assumption.

Proof:

$$\begin{aligned} \hat{r}(x) - \hat{r}(y) &= \sum_{i \in I} \lambda_i (p(x, X_i) - p(y, X_i)) \leq \\ &\left(\sum_{i \in I} \lambda_i \right) (1/2 \sqrt{1 - p(x, y)} + (1 - p(x, y))) \end{aligned}$$

by Theorem 17. ■

13 Conclusion and future work

This paper has proposed SP-structures as a generalization for the structure of one-dimensional subspaces in Hilbert spaces. A novel notion of observables in an SP-structure generalizes self-adjoint, i.e., Hermitian, operators. A physical system should be viewed as an SP-structure together with a set of observables. To be acceptable, the set of observables must be rich enough to justify the similarity measure p of the SP-structure. The value of $p(x, y)$ must be justified by one of the observables at hand. Defining precisely and studying such sets of observables is probably the next step.

Acknowledgements

I want to thank Jean-Marc Lévy-Leblond for his continuing help in straightening me out on Quantum Physics. I also benefitted from Lev Vaidman's help: thanks.

References

- [1] Daniel Lehmann. A presentation of quantum logic based on an "and then" connective. *Journal of Logic and Computation*, 2007. to appear, DOI: 10.1093/logcom/exm054.
- [2] Daniel Lehmann. Quantic superpositions and the geometry of complex hilbert spaces. *International Journal of Theoretical Physics*, to appear.
- [3] Daniel Lehmann, Kurt Engesser, and Dov M. Gabbay. Algebras of measurements: the logical structure of quantum mechanics. *International Journal of Theoretical Physics*, 45(4):698–723, April 2006. DOI 10.1007/s10773-006-9062-y.
- [4] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer, Dordrecht, The Netherlands, 1995.
- [5] Hassler Whitney. On the abstract properties of linear dependence. *American Journal of Mathematics*, 57:509–533, 1935.

A Classical SP-structures

Let Ω be an arbitrary set and let $p(x, y) = 1$ iff $x = y$ and $p(x, y) = 0$ otherwise. We have: $x \sim y$ iff $x = y$. Symmetry is satisfied since equality is symmetric. Non-negativity is satisfied by definition. Two states x and y are orthogonal iff they are different. Any subset A of Ω is an ortho-set and $\bar{A} = A$. Boundedness is satisfied since $p(x, A) = 1$ iff $x \in A$ and $p(x, A) = 0$ otherwise. Suppose, now, that $p(x, A) < 1$. Then, $p(x, A) = 0$ and $p(x, A) + p(x, x) = 1$. One easily sees that x is a suitable y , showing that x is a suitable y in the definition of the property of O-Projection. There is only one basis for Ω : Ω itself. The property of Factorization is easily established. Subspaces (i.e., subsets) X and Y are orthogonal iff their intersection is empty. Orthogonal sum is set union. If $X \perp Y$, $a, b \in X \cup Y$, $\alpha_{X,Y}(a, b) = 0$ in all cases and therefore Inequality is satisfied. The quantity $\rho_{X,Y}(a, b)$ is also always equal to zero.

B Hilbert SP-structures

Assume \mathcal{H} is a Hilbert space on the complex field, Ω is the set of all unit vectors of \mathcal{H} and p is real scalar product: $p(\vec{x}, \vec{y}) = |\langle \vec{x}, \vec{y} \rangle|^2$.

B.1 First properties

We have: $\vec{x} \sim \vec{y}$ iff there is a phase factor such that $\vec{y} = e^{i\varphi} \vec{x}$, by Cauchy-Schwarz. Symmetry is satisfied since $\langle \vec{y}, \vec{x} \rangle = \overline{\langle \vec{x}, \vec{y} \rangle}$. Non-negativity is satisfied by definition. Orthogonality has its usual meaning in Hilbert spaces. An ortho-set is an orthonormal set of vectors. Boundedness follows from the existence of an orthonormal basis for \mathcal{H} that extends any orthonormal set of vectors and the fact that $\vec{x} = \sum_{\vec{b} \in B} \langle \vec{x}, \vec{b} \rangle \vec{b}$. Basis has its usual meaning in Hilbert spaces. Subspaces are *closed* subspaces. Let X be any subspace and assume that \vec{x} is a unit vector. The quantity $p(\vec{x}, X)$ is the square of the norm of the projection of \vec{x} on the subspace X . Let us show that the property of O-Projection is satisfied. If $p(\vec{x}, X) < 1$, the projection of \vec{x} on the subspace X^\perp orthogonal to X , call it \vec{w} , is not null. Let $\vec{y} = \vec{w} / \|\vec{w}\|$. Note that $\|\vec{w}\|^2 = p(x, y)$. Therefore $1 = \|\vec{x}\|^2 = p(x, X) + p(x, y)$. Then \vec{y} is a unit vector that satisfies the properties of O-Projection. Let us check Factorization. Assume $\vec{x}, \vec{y}, \vec{z}$ are unit vectors and X is a subspace. Assume $\vec{y}, \vec{z} \in X$ and $p(\vec{x}, \vec{y}) = p(\vec{x}, X)$. If $\vec{x} \perp X$ then $\vec{x} \perp \vec{z}$ and Factorization is satisfied. Otherwise, let \vec{w} be the non-null projection of \vec{x} on X and let $\vec{v} = \vec{w} / \|\vec{w}\|$. Then $\langle \vec{x}, \vec{z} \rangle = \langle \vec{w}, \vec{z} \rangle = \langle \vec{v}, \vec{z} \rangle \|\vec{w}\|$. Therefore $p(\vec{x}, \vec{z}) = p(\vec{v}, \vec{z}) p(\vec{x}, \vec{v})$.

B.2 Inequality and phases

Let X and Y be orthogonal subspaces and \vec{a}, \vec{b} be unit vectors in their orthogonal sum $X \oplus Y$. Let $\vec{a}_X, \vec{b}_X, \vec{a}_Y$ and \vec{b}_Y be their projections on X and Y respectively. Assume, first, that $\rho_{X,Y}(\vec{a}, \vec{b}) = 0$. Without loss of generality assume that $\vec{a}_Y = \vec{0}$. Then we have $\vec{a} \in X$. If $\vec{b}_X = \vec{0}$ one easily checks that $\alpha_{X,Y}(\vec{a}, \vec{b}) = 0$. If $\vec{b}_X \neq \vec{0}$ one has

$$\alpha_{X,Y}(\vec{a}, \vec{b}) = p(\vec{a}, \vec{b}) - p(\vec{b}, X) p(\vec{a}, t(b, X)) = 0$$

by Factorization.

Assume, now that $\rho_{X,Y}(\vec{a}, \vec{b}) > 0$. By projection the vectors \vec{a} and \vec{b} on X and Y respectively we have $\vec{a} = \vec{a}_X + \vec{a}_Y$ and $\vec{b} = \vec{b}_X + \vec{b}_Y$. But $\|\vec{a}_X\| =$

$\sqrt{p(\vec{a}, X)}$ and similarly for all four terms. Therefore the unit-vector $t(\vec{a}, X)$ is $\frac{\vec{a}_X}{\sqrt{p(\vec{a}, X)}}$ and similarly for the other terms. Therefore

$$\begin{aligned}\langle \vec{a}, \vec{b} \rangle &= \langle \vec{a}_X, \vec{b}_X \rangle + \langle \vec{a}_Y, \vec{b}_Y \rangle = \\ &\sqrt{p(\vec{a}, X) p(\vec{b}, X)} \langle t(\vec{a}, X), t(\vec{b}, X) \rangle + \sqrt{p(\vec{a}, Y) p(\vec{b}, Y)} \langle t(\vec{a}, Y), t(\vec{b}, Y) \rangle \\ \text{Let } \langle t(\vec{a}, X), t(\vec{b}, X) \rangle &= r e^{i\theta} \text{ and } \langle t(\vec{a}, Y), t(\vec{b}, Y) \rangle = s e^{i\psi}. \text{ We have}\end{aligned}$$

$$\begin{aligned}p(\vec{a}, \vec{b}) &= \\ | \langle \sqrt{p(\vec{a}, X) p(\vec{b}, X)} t(\vec{a}, X), t(\vec{b}, X) \rangle + \langle \sqrt{p(\vec{a}, Y) p(\vec{b}, Y)} t(\vec{a}, Y), t(\vec{b}, Y) \rangle |^2 &= \\ p(\vec{a}, X) p(\vec{b}, X) r^2 + p(\vec{a}, Y) p(\vec{b}, Y) s^2 + & \\ 2 \sqrt{p(\vec{a}, X) p(\vec{b}, X) p(\vec{a}, Y) p(\vec{b}, Y)} r s \cos(\psi - \theta). &\end{aligned}$$

But $r^2 = p(t(\vec{a}, X), t(\vec{b}, X))$ and $s^2 = p(t(\vec{a}, Y), t(\vec{b}, Y))$. Therefore $\alpha_{X,Y}(\vec{a}, \vec{b}) = \rho_{X,Y}(\vec{a}, \vec{b}) \cos(\psi - \theta)$. We have proved Inequality.

C Tight bounds in Theorem 17

Let \vec{u} and \vec{v} be orthogonal and let: $\vec{x} = \sqrt{r} \vec{u} + \sqrt{1-r} \vec{v}$ and $\vec{y} = \sqrt{r-\epsilon} \vec{u} + \sqrt{1-r+\epsilon} e^{i\delta} \vec{v}$, for $r \in [0, 1]$ and ϵ and δ close to zero, and $0 < r < 1$.

We have

$$\begin{aligned}p(\vec{x}, \vec{y}) &= r(r-\epsilon) + (1-r)(1-r+\epsilon) + 2\sqrt{r(r-\epsilon)(1-r)(1-r+\epsilon)} \cos(\delta) = \\ r^2(1-\epsilon/r) + (1-r)^2(1+\epsilon/(1-r)) + 2r(1-r) \sqrt{(1-\epsilon/r)(1+\epsilon/(1-r))} \cos(\delta) &= \\ r^2(1-\epsilon/r) + (1-r)^2(1+\epsilon/(1-r)) + 2r(1-r) \sqrt{(1-\epsilon(1-2r+\epsilon)/(r(1-r)))} \cos(\delta) &\end{aligned}$$

We develop to the second order in ϵ and δ .

$$\begin{aligned}p(\vec{x}, \vec{y}) &= r^2 - r\epsilon + (1-r)^2 + (1-r)\epsilon + \\ 2r(1-r)(1-\delta^2/2) - (1-2r+\epsilon)\epsilon(1-\delta^2/2) - 1/4(1-2r)^2/(r(1-r))\epsilon^2 &= \\ r^2 + (1-r)^2 + 2r(1-r) + &\end{aligned}$$

$$\begin{aligned}
& -r\epsilon + (1-r)\epsilon - (1-2r)\epsilon - \\
& r(1-r)\delta^2 - \\
& (1 + 1/4(1-2r)^2/(r(1-r))\epsilon^2 + O(\epsilon\delta^2)
\end{aligned}$$

We conclude that

$$1 - p(x, y) = r(1-r)\delta^2 + \epsilon^2/(r(1-r)) + O(\epsilon\delta^2)$$

Taking $\delta = 0$ and $z = \vec{u}$ Theorem 17 gives $r \leq r - \epsilon + \epsilon + 4\epsilon^2$ showing that the first term $1/2\sqrt{1-p(x, y)}$ in Theorem 17 is tight.

D Hermitian operators and Observables

We are interested in showing that in a Hilbert SP-structure, Hermitian (i.e., self-adjoint) bounded linear operators are *observables* as defined in Definition 8.

Let \mathcal{H} be a separable Hilbert space, Ω the set of rays of \mathcal{H} and let A be a Hermitian operator on \mathcal{H} i.e., A is a bounded linear self-adjoint transformation of \mathcal{H} . We shall define the transformation $r : X \longrightarrow X$ in the following way. Let $\vec{x} \in X$, i.e., \vec{x} is a unit vector of \mathcal{H} . If $A(\vec{x}) \neq \vec{0}$ we define $r(\vec{x}) = A(\vec{x}) / \|A(\vec{x})\|$ and if $A(\vec{x}) = \vec{0}$ we define $r(\vec{x}) = \vec{x}$.

We are going to show that r is an observable in the sense of Definition 8. The set of eigenvalues of A is the set of λ_i 's for $i \in I$. For $i \in I$ the subspace X_i is the set of all unit vectors of the eigensubspace of A corresponding to λ_i . Let \vec{a}, \vec{b} be unit vectors and assume \vec{b} is an eigenvector for λ_i . If $\sum_{j \in I} \lambda_j^2 p(a, X_j) = 0$, then \vec{a} is an eigenvector for the eigenvalue 0 and therefore $r(\vec{a}) = \vec{a}$ and $p(r(\vec{a}), \vec{b}) = p(\vec{a}, \vec{b})$ for any unit vector \vec{b} . If $\sum_{j \in I} \lambda_j^2 p(a, X_j) > 0$, then

$$\begin{aligned}
p(r(\vec{a}), \vec{b}) &= \left| \frac{\langle A(\vec{a}), \vec{b} \rangle}{\|A(\vec{a})\|} \right|^2 = \left| \frac{\langle \vec{a}, A(\vec{b}) \rangle}{\|A(\vec{a})\|} \right|^2 = \left| \frac{\langle \vec{a}, \lambda_i \vec{b} \rangle}{\|A(\vec{a})\|} \right|^2 = \\
& \frac{\lambda_i^2 p(\vec{a}, \vec{b})}{\sum_{j \in I} \lambda_j^2 p(\vec{a}, X_j)}.
\end{aligned}$$

Let us show, now, that item 4 of Definition 8 holds. Let $j, k \in I$ such that $j \neq k$ and $\vec{a}, \vec{b} \in \Omega$ such that neither \vec{a} or \vec{b} are orthogonal to either X_j or X_k . Let $Z = X_j \oplus X_k$. The projections $t(\vec{a}, Z)$, $t(\vec{b}, Z)$ and $t(r(\vec{a}), Z)$

exist (by Theorem 21 item 3). Let $t(\vec{a}, Z) = \sqrt{r} \vec{a}_1 + \sqrt{1-r} e^{i\varphi} \vec{a}_2$ with $0 < r = p(t(\vec{a}, Z), \vec{a}_1) = p(\vec{a}, \vec{a}_1) / (p(\vec{a}, \vec{a}_1) + p(\vec{a}, \vec{a}_2)) < 1$. Similarly, let $t(\vec{b}, Z) = \sqrt{s} \vec{b}_1 + \sqrt{1-s} e^{i\theta} \vec{b}_2$. By the analysis of Appendix B.2 we have:

$$\alpha_{X_j, X_k}(\vec{a}, \vec{b}) = \rho_{X_j, X_k}(\vec{a}, \vec{b}) \cos(\theta - \varphi).$$

Assume, first, that both λ_j and λ_k are strictly positive. We have $A(\vec{a}) = \lambda_j \sqrt{r} \vec{a}_1 + \lambda_k \sqrt{1-r} e^{i\varphi} \vec{a}_2$ is a non-null vector and its norm is $R = \sqrt{\lambda_j^2 r + \lambda_k^2 (1-r)}$. Therefore $t(r(a), Z) = A(\vec{a}) / R$ and

$$(3) \quad \alpha_{X_j, X_k}(r(\vec{a}), \vec{b}) = \rho_{X_j, X_k}(r(\vec{a}), \vec{b}) \cos(\theta - \varphi).$$

We see that $\omega_{X_j, X_k}(r(\vec{a}), \vec{b}) = \omega_{X_j, X_k}(\vec{a}, \vec{b})$.

Assume, now that both λ_j and λ_k are strictly negative. Then, the proper form for a non-null vector of $t(r(\vec{a}), Z)$ is now $(-\lambda_j) \sqrt{r} \vec{a}_1 + (-\lambda_k) \sqrt{1-r} e^{i\varphi} \vec{a}_2$. and we also obtain Equation 3.

Assume, now that $\lambda_j > 0$ and $\lambda_k < 0$. Then, the proper form for a non-null vector of $t(r(\vec{a}), Z)$ is $\lambda_j \sqrt{r} \vec{a}_1 + (-\lambda_k) \sqrt{1-r} e^{i\varphi+\pi} \vec{a}_2$. Then

$$\alpha_{X_j, X_k}(r(\vec{a}), \vec{b}) = \rho_{X_j, X_k}(r(\vec{a}), \vec{b}) \cos(\theta - \varphi + \pi).$$

We obtain $\omega_{X_j, X_k}(r(\vec{a}), \vec{b}) = -\omega_{X_j, X_k}(\vec{a}, \vec{b})$.

Similarly, the same obtains if $\lambda_j < 0$ and $\lambda_k > 0$.